

## Reading the face of a leader: women with low facial masculinity are perceived as competitive

Article (Published Version)

Silberzahn, Raphael and Menges, Jochen (2016) Reading the face of a leader: women with low facial masculinity are perceived as competitive. *Academy of Management Discoveries*, 2 (3). pp. 272-289. ISSN 2168-1007

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/73376/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

### **Copyright and reuse:**

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

# READING THE FACE OF A LEADER: WOMEN WITH LOW FACIAL MASCULINITY ARE PERCEIVED AS COMPETITIVE

RAPHAEL SILBERZAHN  
University of Cambridge - Judge Business School

JOCHEN MENGES  
WHU - Otto Beisheim School of Management



In competitive settings, people prefer leaders with masculine faces. But is facial masculinity a trait that is similarly desired in men and women leaders? Across three studies, we discovered that people indeed prefer men and women leaders who have faces with masculine traits. But surprisingly, we find that people also prefer women with low facial masculinity as leaders in competitive contexts (Study 1). Our findings indicate that low facial masculinity in women, but not in men is perceived to indicate competitiveness (Study 2). Thus, in contrast to men, women leaders who rate high in facial masculinity as well as those low in facial masculinity are both selected as leaders in competitive contexts. Indeed, among CEOs of S&P 500 companies, we find a greater range of facial masculinity among women CEOs than among men CEOs (Study 3). Our results suggest that traits of facial masculinity in men and women are interpreted differently. Low facial masculinity in women is linked to competitiveness and not only to cooperativeness as suggested by prior research.

## Editor's Comment

The paper by Silberzahn and Menges provides important insights into the effects of masculine and feminine facial features. Do leaders in competitive situations profit from masculine or feminine features, and does this vary across men and women? Based on their intuition and a few available data points, the authors expected mainstream ideas to hold for men but not for women, where those ideas support masculinity for leaders embedded in strongly competitive situations. The authors were right—among men, those with masculine features were favored as leaders and indeed were seen as more competitive, but among women those with strongly masculine features *and* those with strongly feminine features were favored and were rated as more competitive. Having shared these findings with several executive audiences, this editor can tell you that practicing managers are quite intrigued by the implications. For scholarship, the findings reported in the paper have implications for several research streams and theories.

C. Chet Miller, Action Editor

## INTRODUCTION

Yahoo's Marissa Mayer, Hewlett Packard's Meg Whitman, and Facebook's Sheryl Sandberg have in particular three things in common: They are all top-level leaders in highly competitive companies, they are all women, and none of them look particularly masculine. Yet, previous research studies have found that people prefer men and women leaders with masculine faces, particularly in competitive settings (Little, Burriss, Jones, & Roberts, 2007; Little, 2014; Re, DeBruine, Jones, & Perrett, 2013; Spisak, Dekker, Krüger, & van Vugt, 2012; Spisak, Homan, Grabo, & Van Vugt, 2012). Are Mayer, Whitman, and Sandberg then just exceptions to the rule, or is there something more systematic to the success of nonmasculine-looking women in the corporate world?

Women and men whose faces appear masculine are more likely to attain leadership positions, as high facial masculinity signals relevant personality traits such as dominance and agency (Re et al., 2013; Spisak, Dekker, et al., 2012). These traits, according to leadership theory (Hogue & Lord, 2007) and gender research (Eagly & Karau, 2002), are seen as desirable for leaders. Faces of potential leaders are read for cues to masculinity, and these cues appear to be even more influential than the person's biological sex in influencing leader choice (Spisak, Dekker, et al., 2012). Higher degrees of facial masculinity thus seem to be desired in both men and women leaders.

However, the appearance of men and women is often evaluated differently (Feinberg and Jones, 2005; Rhodes, 2006; Roney, Hanson, Durante, & Maestripieri, 2006; Smith et al., 2006). Women with feminine faces are considered more attractive (Thornhill & Gangestad, 1999), which may help in attaining positions with higher occupational prestige (Jæger, 2011). Results from simulated political elections suggest preferences for women with less masculine faces (Hehman, Carpinella, Johnson, Leitner, & Freeman, 2014). Hence, although facial masculinity appears to benefit men's attainment of leadership positions, for women, the case is less clear-cut. This led us to wonder if facial masculinity affects women and men differently in leadership selections?

In a series of three experimental and field studies, we examined whether facial masculinity affects men and women leaders differently, and why this may be the case. We made three discoveries: First, whereas

men in competitive settings benefit from high levels of facial masculinity, women fare well when they either look particularly masculine or when they do not look masculine at all. Second, in contrast to prior research, we find that among women, low facial masculinity is perceived to signal competitiveness, whereas among men, competitiveness was inferred only from high facial masculinity and not from low facial masculinity. Third, our findings suggest that the differential effects of facial masculinity on leadership preferences manifest in the corporate world. We discover that in S&P 500 companies, a greater range of facial masculinity is present among women CEOs compared to men CEOs.

The association between low facial masculinity and competitiveness in women, which we detect in our research, adds to and challenges existing theory and research across several literature streams. We extend gender role research by offering a more nuanced understanding of how attributions of competitiveness vary not just between, but also within gender (Eagly & Karau, 2002). Our discoveries are relevant for the further development of theories concerning leadership prototypes (Hogue & Lord, 2007): Because both women with high and low masculinity are perceived favorably as potential leaders, there may be less of a disadvantage for women based on people's implicit leadership theories than previously assumed. Our research extends and refines biosocial models of leadership (Spisak, Dekker, et al., 2012). It encourages further research linking leadership choice and attributions about facial masculinity with hormonal effects on facial appearance and behavior (Stanton & Schultheiss, 2007, 2009). And, our findings are of interest to personality researchers who have long wondered about the origin and validity of personality attributions made from facial appearance (Penton-Voak et al., 2006; Zebrowitz & Montepare, 2008).

## FACIAL MASCULINITY AND LEADERSHIP PREFERENCES

When choosing a leader, people are inevitably influenced by others' facial appearance. People process facial information in a split second and form instantaneous impressions (Willis & Todorov, 2006). This process is reinforced by biosocial models of leadership suggesting that a likely leader's behavior

Author's voice:

What additional findings surprised or motivated you?



Author's voice:

What motivated you personally to undertake this research?



is inferred from masculine facial features (Spisak, Dekker, et al., 2012). Facial features are shaped in the womb and during later life by sex hormones (Boothroyd & Perrett, 2006; Lutchmaya, BaronCohen, Raggatt, Knickmeyer, & Manning, 2004). The male sex hormone testosterone makes faces more masculine (Lefevre, Lewis, Perrett, & Penke, 2013), and the female sex hormone estrogen makes faces appear more feminine (Smith et al., 2006). But in addition to influencing facial appearance, hormones also affect behavior. Testosterone is associated with status striving and dominance in both men (Pound, Penton-Voak, & Surridge, 2009; Stanton & Schultheiss, 2009) and women (Jiménez, Aguilar, & AlveroCruz, 2012). Thus, biosocial models of leadership posit that over the course of evolution, masculine facial features have become linked in people's brains with positive inferences concerning leadership behavior:

The more masculine a face looks, the more likely the person is believed to engage in behaviors that are typically expected from leaders.

Biosocial models of leadership suggest that people's perceptions of faces and leadership preferences are linked through implicit leadership theories. Implicit leadership theories operate outside of awareness and facilitate people's instantaneous assessment of the leadership qualities of another person (Lord & Dinh, 2014). People use these prototypical expectations to automatically assess whether the leader matches their prototype. If there is a match, then the leader is considered more favorably than if there is no match (Lord, Foti, & De Vader, 1984).

Masculine faces are assumed to match with leadership prototypes because facial stimuli activate schemas. Schemas are organized mental representations of life experiences that facilitate the processing of new information (Baldwin, 1992), allowing individuals to make inferences about a person based on minimal information (Neisser, 1976). For example, so-called baby faces activate a schema that attributes incompetency and a lack of maturity to a person, even in the absence of relevant evidence (Zebrowitz & Montepare, 2008). It is the activated schema that fills in the missing information. Similarly, masculine faces activate a schema that evokes dominance and strength (Little, 2014), traits that are characteristic of prototypical leaders (Lord & Dinh, 2014). Masculine

faces are likely associated with dominance and power, because in many animal species, including primates (De Waal, 1982) and humans (Murray & Schmitz, 2011), facial masculinity is linked to greater physical size and strength, body features which in many species determine social status and rank (Henrich & Gil-White, 2001; Little et al., 2007; Re et al., 2013). This pattern was also evident in a study of elementary school children who ascribed strength, resourcefulness, independence and, indeed, leadership to those girls whom they categorized as rather masculine (Hemmer & Kleiber, 1981).

The association between masculinity and leadership is captured in the "think manager, think male" phenomenon (Schein, 1973, 2001). Although preferences are changing, the prototypical expectations for leaders still favor men (Newport & Wilke, 2013). In line with these preferences, masculine facial features are seen as fitting the admittedly gender-biased leadership prototype (Hogue & Lord, 2007).

But the link between facial masculinity and leadership preferences also depends on the context of leadership. Implicit leadership theories are not fixed, but rather adjust according to a given context such that different contexts give rise to different prototypical expectations about leaders (Brands, Menges, & Kilduff, 2015). The biosocial model of leadership posits the contingency that preferences for high facial masculinity apply particularly in situations that involve greater conflict, competition, and aggression, whereas preferences for comparatively lower facial masculinity apply in more peaceful, cooperative settings (Little et al., 2007; Spisak, Dekker, et al., 2012). Such settings may cater to a communal, nurturing leadership style aligned with the female gender stereotype, whereas an agentic dominant leadership style appears suited to competitive settings and is aligned with the male gender stereotype (Eagly & Karau, 2002).

Thus, situational demands activate different leadership prototypes, and people consequently prefer leaders whom they expect to have characteristics suited to that situation (Little, 2014). The contingency advanced by the biosocial model of leadership (Spisak, Homan, et al., 2012) therefore suggest a person-in-situation schema (Brands et al., 2015) that favors masculine-looking leaders in competitive contexts. In such contexts, both men and women who lack facial masculinity may fail to meet people's

Author's voice:

Was there anything that surprised you leadership selections?



Author's voice:

The link between facial masculinity and leadership preferences



assumptions, as suggested by leadership categorization accounts of implicit leadership theories (Hogue & Lord, 2007), gender research and, particularly, by role congruity theory (Eagly & Karau, 2002).

Yet, in three studies, we discovered a more nuanced pattern of linkages between facial masculinity and leadership preferences for women leaders. In Study 1, we asked individuals to choose potential leaders from a set of candidates with a range of modified degrees of facial masculinity. We manipulated leader gender (male/female) and leadership context (competitive/collaborative), and found that preferences for men and women leaders differ. Women scoring high on masculinity but also those with low facial masculinity were preferred in competitive settings. In Study 2, we examined why these effects occur. With a categorization task, we studied what attributes and characteristics people ascribe based on variations in facial masculinity. We found that women leaders who scored high *as well as* those who scored low in facial masculinity were attributed competitiveness, whereas for men leaders, this was the case only for those high in facial masculinity. If these effects manifest not only in laboratory settings, but also in field settings, we would expect to find greater range of facial masculinity among women than among men leaders. Indeed, in Study 3, this was confirmed using a sample of male and female CEOs of S&P 500 companies. In summary, the findings across the three studies challenge conclusions in existing research on facial masculinity and leadership preferences, and instead suggest that women and men are affected differently by facial masculinity.

## STUDY 1

We examined whether preferences for facial masculinity differ for men and women leaders in competitive and collaborative contexts. We devised a  $2 \times 2$  between-subjects experiment in which two factors—leader gender (man versus woman) and context (competitive versus collaborative)—were varied. Prior research suggests that in competitive environments, people would prefer leaders with higher levels of facial masculinity, and that such preferences would be similar for men and women (Spisak, Dekker, et al., 2012). In contrast, our thinking and some related research studies (Hehman et al., 2014) suggest that people might choose women with faces low in masculinity as leaders in competitive environments.

## Participants and Procedures

The sample consisted of 187 American adults (44 percent women) from different ethnicities (67 percent White, 14 percent Black, 10 percent Asian, 4 percent Latino, and 4 percent other). All participants resided in the United States, and they averaged 33 years of age (standard deviation [SD] = 12.04) and 12 years of work experience (SD = 11.35). Participants were recruited from Microworkers, an online panel.

Participants accessed our website and, after agreeing to take part in this research, were randomly assigned to one of four conditions. Participants would initially read that a company was either competing heavily or collaborating well, and were then asked to make several leadership choices by selecting a suitable leader from a series of images showing faces of women or men. Each image featured the same person's face but with varied degrees of masculinity.

**Manipulation 1: Leader gender.** Participants in the men leader condition were shown a series of images of men, and participants in the women leader condition were shown a series of images of women.

**Manipulation 2: Leadership context.** Participants in the competition condition read a scenario that described a company that “has many rivals and competes heavily,” and participants in the collaboration condition read about a company that “has many partners and collaborates well.” The scenarios were taken from prior research (Little et al., 2007) and adapted to a corporate context.<sup>1</sup> Each scenario was reinforced by adding to the subsequent questions concerning leader choice the phrases “. . . given that your company collaborates well with other companies” or “. . . given that your company competes heavily with other companies.”

**Dependent variable: Leader choice.** We asked candidates four questions concerning whom they would (1) “prefer for your company,” (2) “hire for your company,” (3) find “most effective,” and (4) believe is “most likely to achieve your company's goal.” Based on sufficient internal consistency ( $\alpha = 0.79$ ), we combined these four items in an overall measure of leader choice. Participants indicated their choice for each item by selecting one image from a series of five images of men (in the men condition) or women (in the women condition). The

<sup>1</sup> We pretested whether the manipulation would lead people to perceive a company as either competitive or collaborative in a sample of 77 American adults. After reading each scenario, participants reported significantly different ratings concerning competition versus collaboration,  $t(76) = 2.74$ ,  $p = .02$ . Thus, the manipulation proved effective.

Author's voice:

How did the paper evolve and change as you worked on it?





five images<sup>2</sup> showed the same basic male (female) face of a Caucasian adult. The images featured different degrees of masculinity, forming a scale from low masculinity (1) to high masculinity (5), with each number indicating a 50 percent decrease or increase in masculinity relative to the original composite: 1 = 100 percent masculinity, 2 = 50 percent masculinity, 3 = original composite, 4 = +50 percent masculinity, and 5 = +100 percent masculinity. All five images were displayed in one row after each of the four questions mentioned above, but the order of presentation of the images was randomized to avoid displaying an obviously linear masculinity scale. Based on participants' responses, we calculated two scores. First, we calculated participants' *overall leader masculinity preference* by averaging the responses on the 1–5 masculinity scale across the four leadership choice questions. Second, we recorded participants' *specific leader masculinity preference* for each of the five leader images, counting for how many of the four questions participants selected each of the leader images. When a participant chose a specific leader image, the image received a score of 1. Across the four questions, each leader image thus received scores ranging from 0 (the participant chose the image for none of the leader choice questions) to 4 (the participant chose the image for all of the four questions) from each participant. We then calculated a mean value for each leader image, indicating how often participants on average chose this image. (For ease of interpretation, we also report the relative percentages, again indicating how often participants on average chose the leader image.)

### Data Quality Check

The data were collected through an online panel, thus outside a supervised laboratory. Although this approach is increasingly common (Bohannon, 2011) and the data thus collected tend to be of similar or better quality than the data collected in laboratory studies (Buhrmester et al., 2011), there is a need to ensure that participants followed the instructions and provided meaningful answers. In line with other recent organizational research using online panels (Brands et al., 2015), we thus checked the quality of the data prior to our analyses, based on four criteria. Specifically, we

<sup>2</sup> The images were taken from a previous study with permission of the corresponding author (Re et al., 2013). The images were created with the software Psychomorph following established procedures (Perrett et al., 1998). Differences in masculinity reflect relative distances between facial landmarks of the man and woman's faces.

**TABLE 1**  
**Study 1: Overall Leader Masculinity Preferences in Collaborative versus Competitive Contexts**

Leader Gender	Leadership Context	Mean	SD	N
Woman leader	Collaboration	3.03	0.84	44
	Competition	3.14	1.26	52
	Total	3.09	1.08	96
Man leader	Collaboration	3.52	0.91	42
	Competition	3.97	0.74	49
	Total	3.76	0.85	91
Total	Collaboration	3.27	0.91	86
	Competition	3.54	1.12	101
	Total	3.42	1.03	187

established that participants' responses would be excluded if (1) participants' response time was longer than 2 SD above the mean, as extended response times may imply that distractions occurred during participation; (2) participants had the same IP address, as this would indicate repeated participation; (3) a participant's response pattern indicated a lack of attention to the questionnaire items (e.g., all images were selected from just one side of the screen); or (4) participants incorrectly answered an attention check question<sup>3</sup> that was included in the assessment of demographics toward the end of the questionnaire. Based on these four criteria, none of the participants had to be excluded.

### Results

We were interested in whether preferences for facial masculinity differ for men and women leaders in competitive and collaborative contexts. Thus we began our analyses with an analysis of variance (ANOVA), entering leader gender (women leaders versus men leaders) and leadership context (competition versus collaboration) as fixed factors, and overall leader masculinity preference as the dependent variable. The means and SDs for each condition are listed in Table 1. Participants preferred high facial masculinity more for men leaders than for women leaders, as indicated by a significant main effect of leader gender,  $F(1, 183) = 21.34, p < .01$ . Participants also preferred high facial masculinity more for competitive contexts than for collaborative contexts, as indicated by a significant main effect of leadership context,  $F(1, 183) = 3.83, p = .05$ . Participants' preferences were not dependent on the specific configuration of leader gender

<sup>3</sup> The multiple choice question asked participants for how many hours on average they slept, and offered as answer options: 1 = 0 hours, 2 = 410 hours, 3 = 24 hours. If participants report 1 or 3, they show a lack of attention to the question.

**TABLE 2**  
**Study 1: The Effect of Men's Facial Masculinity on Leader Choice in Collaborative versus Competitive Contexts**

Masculinity	1 (−100%)	2 (−50%)	3 (original)	4 (+50%)	5 (+100%)
Leadership Context	Leader Choice				
Collaboration					
Percentage	8	15	18	37	23
Mean	0.31	0.60	0.71	1.48	0.90
SD	0.60	0.94	1.04	1.29	1.14
Competition					
Percentage	3	6	20	34	37
Mean	0.12	0.22	0.80	1.37	1.49
SD	0.44	0.65	1.04	1.25	1.45

*Note.* This table shows how often a particular leader was selected in each condition. Percentages, means, and SDs indicate how many times on average each leader image was selected for the four leader choice questions.

Number of collaboration = 42 participants.

Number of competition = 49 participants.

and leadership context: There was no significant interaction effect,  $F(1, 183) = 1.48, p = .23$ . We examined whether the inclusion of participant gender would affect results in an analysis of covariance (ANCOVA). This was not the case.

The results of these analyses are in line with prior theory and research concerning facial masculinity and leadership choice. However, we suspected that these analyses might disguise preferences for women leaders with low facial masculinity in competitive environments. Indeed, a visual inspection of the distribution of specific leader masculinity preferences suggested that participants' leader choices followed different patterns, depending on leader gender. To test whether the response distributions for leader choices for men and women were significantly different, we applied a Kolmogorov–Smirnov test. This nonparametric test is commonly used to check whether data are normally distributed by testing the difference between a sample distribution and a standard normal distribution. But this test can also

be used to examine whether any two distributions are significantly different from one another. We examined whether the distribution of leader choices varied by leader gender in collaborative and competitive contexts. The results indicated that choices for men leaders follow a different distribution than choices for women leaders, both in competitive ( $p < .01$ ) and collaborative settings ( $p < .01$ ).

To examine whether the distributions differ at the low end or the high end of facial masculinity, we plotted the response frequencies for men and women leaders for both the collaborative and the competitive condition in Tables 2 and 3 (see also Figure 1). For men, Table 1 shows the familiar pattern of a preference for higher masculinity in competitive settings. Across the four leader choice questions, participants selected the image of a man with very high facial masculinity (+100 percent) more often in competitive settings ( $M = 1.48$ ; 37 percent) than in collaborative settings ( $M = 0.90$ ; 23 percent). The image of a man with very low facial

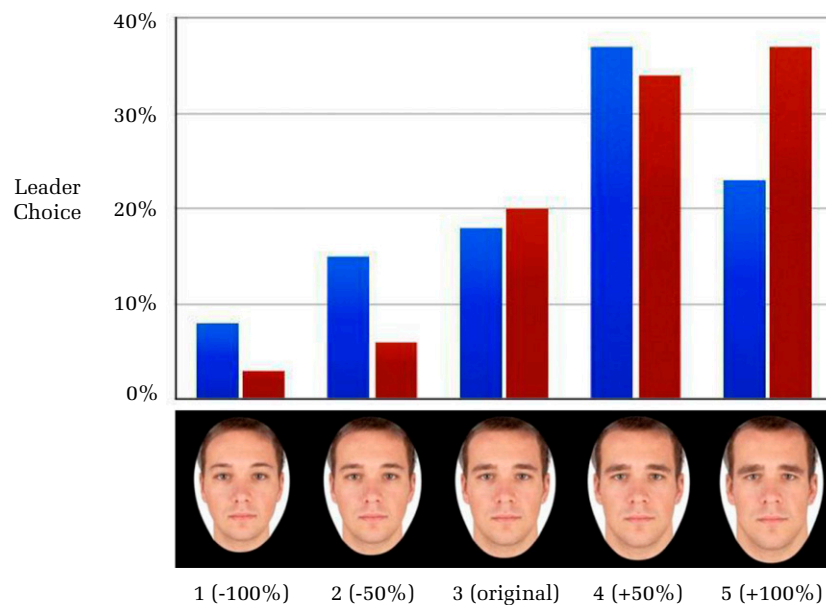
**TABLE 3**  
**Study 1: The Effect of Women's Facial Masculinity on Leader Choice in Collaborative versus Competitive Contexts**

Masculinity	1 (−100%)	2 (−50%)	3 (original)	4 (+50%)	5 (+100%)
Leadership Context	Leader Choice				
Collaboration					
Percentage	7	33	27	16	17
Mean	0.27	1.32	1.09	0.64	0.68
SD	0.45	1.23	1.12	1.04	1.18
Competition					
Percentage	20	18	19	12	30
Mean	0.81	0.73	0.77	0.48	1.21
SD	1.37	1.05	1.08	0.70	1.40

*Note.* This table shows how often a particular leader was selected in each condition. Percentages, means, and SDs indicate how many times on average each image was selected for the four leader choice questions.

Number of collaboration = 44 participants.

Number of competition = 52 participants.

FIGURE 1<sup>a</sup>

Note. Blue refers to the collaborative context and red refers to the competitive context. Values are also shown in Table 2.

<sup>a</sup>Image produced by Re et al., 2013, and is reproduced with permission of the corresponding author.

masculinity was rarely chosen, and less often in the competitive settings ( $M = 0.12$ ; 3 percent) than in collaborative settings ( $M = 0.31$ ; 8 percent). We conducted  $t$  tests to examine whether these results were statistically significant. Men with high facial masculinity (+100 percent) were selected more often as leaders in the competitive than in the collaborative scenario [ $M = 1.49$  versus  $M = 0.90$ ,  $t(89) = 2.122$ ,  $p = .037$ ]. Men with low facial masculinity (100 percent) were not selected more often; to the contrary, they tended to be selected less often in the competition than in the collaboration scenario, with results trending toward significance [ $M = 0.31$  versus  $M = 0.12$ ,  $t(89) = 1.705$ ,  $p = .092$ ].

The pattern for women leaders looks remarkably different however. Figure 2 shows that the image of a woman very high in facial masculinity was chosen more frequently across the four leader choice questions in the competitive context ( $M = 1.21$ ; 30 percent) compared to the collaborative context ( $M = 0.68$ ; 17 percent); and the image of a woman very low in facial masculinity was also chosen more frequently in the competitive context ( $M = 0.81$ ; 20 percent) compared to the collaborative context ( $M = 0.27$ ; 7 percent).  $t$  tests that compared how often on average participants selected women with high and low facial masculinity in competitive as opposed to collaborative environments confirmed the pattern of results. Women leaders with high facial masculinity were selected more often in the competitive than in the

collaborative condition [ $M = 1.21$  versus  $M = 0.68$ ,  $t(94) = 1.981$ ,  $p = .047$ ], and women with low facial masculinity were *also* selected more often in the competition than in the collaboration scenario [ $M = 0.81$  versus  $M = 0.27$ ,  $t(94) = 2.473$ ,  $p = .015$ ].<sup>4</sup>

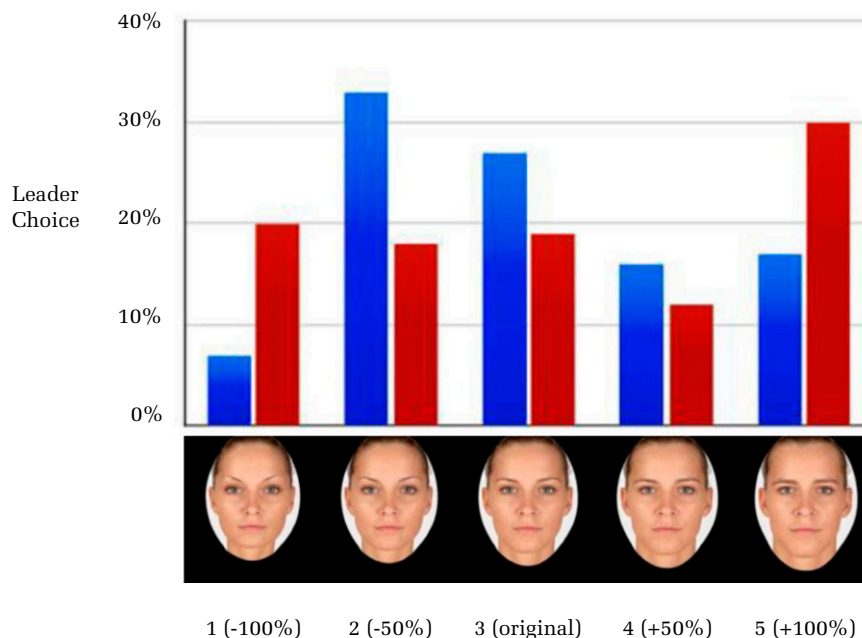
Thus the difference between preferences for facial masculinity between men and women leaders manifests at the low end of the facial masculinity scale. When choosing a man leader for a competitive company environment, participants tended to prefer faces that were very high in masculinity, but when choosing a woman leader for such an environment, participants tended to prefer faces that were either very high in masculinity or very low in masculinity.

## Discussion

The findings from this study indicate that people systematically select men and women for leadership positions based on facial features. However,

<sup>4</sup> As robustness checks, we conducted four regression analyses with participants' gender as a control variable. Results indicate that women leaders with high ( $b = 0.54$ ,  $p = .05$ ) and with low facial masculinity ( $b = 0.51$ ,  $p = .02$ ) were selected more often in the competitive than in the collaborative scenario. Men leaders with high ( $b = 0.57$ ,  $p = .05$ ) but not those with low masculinity ( $b = 0.16$ ,  $p = .04$ ) were selected more often in the competitive than in the collaborative scenario.



FIGURE 2<sup>a</sup>

*Note.* Blue refers to the collaborative context and red refers to the competitive context. Values are also shown in Table 3.

<sup>a</sup>Image produced by Re et al., 2013, and is reproduced with permission of the corresponding author.

participants' preferences for facial masculinity follow different patterns for men and women leaders, which stands in contrast to findings from prior research (Spisak, Dekker, et al., 2012). We did indeed find that high masculinity (in men and women) is more preferred in competitive than in collaborative scenarios (Re et al., 2013). Yet, in competitive scenarios, women leaders with low facial masculinity were also preferred more than in collaborative settings. In contrast, for men leaders with low masculinity, this was not the case. This preference for low masculinity among women leaders in a scenario of competition is somewhat surprising, given that low facial masculinity has been linked to perceptions of warmth, cooperativeness, and lower levels of dominance (Perrett et al., 1998).

Why then is it that in this study, women with low facial masculinity are preferred more in competitive than in collaborative settings? Previous research suggests that people choose leaders with facial features that signal traits that are likely beneficial for a particular task (Little, 2014). Given that women with low facial masculinity were more often chosen in competitive settings, we wondered whether women with low facial masculinity are attributed competitiveness. That is what we chose to examine in Study 2.

## STUDY 2

Do inferences from facial masculinity concerning competitiveness and collaborativeness differ for men and women? In Study 2, we addressed this question through a categorization task. In a prestudy, we first created the stimulus material by asking participants to write essays about men and women leaders. From those essays, we extracted statements, or idea units (Chafe, 1980), describing behaviors, characteristics, and experiences of leaders. Next, we asked independent coders to categorize each extracted statement in terms of whether it was indicative of someone who exhibited competitive behavior.

In the main study, we asked participants (differing from those of the prestudy) to assign each statement to a leader with high facial masculinity, or to one with low facial masculinity, or to neither. We wondered if participants would associate statements indicative of competitive behavior to women with both high and low facial masculinity but only to men high in facial masculinity. This would help us better understand findings from Study 1.

### Prestudy

The prestudy served to obtain content-rich stimulus material that captured the behaviors, characteristics,

and experiences that ordinary people associate with successful leaders in competitive businesses. We asked 44 American adults (43 percent women; 85 percent White, 7 percent Asian, 4 percent Latino, and 4 percent Black) recruited through Amazon Mechanical Turk to write a success story about a business leader. Participants averaged 33 years of age ( $SD = 10.19$ ) and 12 years of work experience ( $SD = 10.15$ ).

After agreeing to take part in our research, participants accessed our website and were instructed to write an essay about the success of the CEO of a fictional company that we named ACME Inc. We described the company as a major movie production company, and continued as follows: "ACME is operating in a highly competitive industry, producing expensive movies to be screened in movie theaters all over the world. When the CEO took over 5 years ago, ACME was in trouble. It lagged behind its competitors, it lacked good movie ideas, and its top managers were estranged. Five years later, these problems are history, and ACME has produced a record profit this year. Your article should describe what led to this success." We also showed participants an image with the caption "CEO of ACME Inc." that was taken from Study 1. The image was either a man or a woman, and either high, medium, or low facial masculinity. Thus, we obtained a diverse set of essays about successful leaders, and protagonists were not only men and women leaders with high, but also leaders with medium and low facial masculinity. After viewing the photo and reading the instructions, participants proceeded to the next page, where they were given space to write their essay. Three text boxes gave participants some structure for their text. The first text box was titled "CEO characteristics," the next "Coworkers," and the last "Competitors."

We separated the essays by the gender of the protagonist, according to the image we had shown to participants.<sup>5</sup> By keeping texts about men and women protagonists separate, we retained gender-specific meanings and connotations (e.g., being a varsity cheerleader or captain of the soccer team may have different connotations for men than for women). From all essays, we extracted only those statements that contained general descriptions of the protagonist, such as those describing characteristics (e.g., "is very hands on"), experiences (e.g., "valedictorian of her class for high school and college"), or behavior (e.g., advice giving: "many of his friends come to him for business advice"). We ignored other statements that were context specific (e.g., "[His] strategy has been to move

away from flagship movies and [he] invested time and resources into R&D") or that were unrelated to the protagonist (e.g., "ACME had been spreading itself thin in the past, trying to please all audiences from kids to senior citizens"). After obvious duplicates were removed, this process resulted in 211 statements describing men's behavior, characteristics, and experience, and 250 statements describing women's behavior, characteristics, and experience in a leadership context.

### Coding of Stimulus Material

To code the extent to which the statements captured the central variables of competitive and collaborative qualities, we handed the statements to two independent coders. The focus of this study was on lay perceptions, thus we hired two coders from the general American population through the online platform Odesk (now called Upwork). The coders live in different parts of the United States and offer research support on a freelance basis. One coder was a woman with 16 years of work experience, and the other coder a man with 6 years of work experience. Both were rated highly for the quality of their work on Odesk. The coders were unaware of the purpose of this study. We asked each coder separately whether a given statement described a behavior, characteristic, or experience that was indicative of someone who is good at competing, collaborating, or neither. As categorization was not straightforward for many statements, we decided beforehand that only statements would be retained for which both coders independently agreed on a category. This approach is analogous to scale development, where items that show poor fit are routinely culled (Hinkin, 1995). In our context, coders converged on the same categorization for 126 (60 percent) of the 211 statements about men; and for 167 (67 percent) of the 250 statements about women. Of these statements, 42 were indicative of competition for men, and 62 were indicative of competition for women. The remaining statements were indicative of collaboration (68 for men, 68 for women) or of neither competition nor collaboration (16 for men, 37 for women).

### Main Study

We recruited 252 American adults (53 percent women; 74 percent White, 8 percent Asian, 7 percent Latino, 8 percent Black, and other 3 percent) via the online panel Microworkers. Participants averaged 32 years of age ( $SD = 11.6$ ) and 11 years of work experience ( $SD = 10.15$ ).

After agreeing to take part in this research, participants were randomly allocated to one of four groups. In each group, participants were shown two images of faces and a list of statements. The images were taken from Study 1, and the statements were taken from the prestudy. In group 1, participants saw

<sup>5</sup> One essay was excluded, because the participant did not correctly recognize the gender of the depicted CEO. Although we had shown the participant the image of a woman, the participant used male pronouns in the essay.

**TABLE 4**  
**Study 2: GEE Regression Results Showing Attributions of Collaborative and Competitive Statements to Men and Women Leaders with Low and High Facial Masculinity**

Leader Gender	Model 1		Model 2		Model 3		Model 4	
	Men Leaders				Women Leaders			
	Low		High		Low		High	
	Masculinity	Odds ratio	95% CI	Odds ratio	95% CI	Odds ratio	95% CI	Odds ratio
Participant gender <sup>a</sup>	0.91	0.81–1.03	1.38**	1.20–1.60	0.94	0.83–1.07	1.07	0.68–1.70
Collaborative attributes	1.04	0.85–1.27	0.50**	0.39–0.65	0.97	0.83–1.14	2.06**	1.40–3.08
Competitive attributes	0.82 <sup>†</sup>	0.66–1.02	1.34*	1.04–1.73	3.39**	2.89–3.97	3.41**	2.44–4.76
Model fit (QICC)	6,021.59		4,583.26		5,824.18		5,381.26	
Number of statements <sup>b</sup>	126		126		167		167	
Number of observations	4,625		3,912		4,543		4,045	

Note. Entries show OR and 95 percent confidence intervals (CI). OD higher (lower) than 1 indicate a positive (negative) effect. The corrected quasi likelihood under Independence Model Criterion (QICC) for model fit is in the “small is better format”).

<sup>a</sup>Woman = 1, Man = 0.

<sup>b</sup>Respondents were assigned to a random subset of 80 statements. For Models 1 and 2, the number of respondents per statement ranged between 40 and 56. For Model 3, the number of respondents per statement ranged between 24 and 64. For Model 4, the number of respondents per statement ranged between 24 and 66.

\* $p < .05$

\*\* $p < .01$

<sup>†</sup> $p < .10$

two male faces: One of the faces was the unmodified original composite, while the comparison image had been modified to depict high masculinity (+100 percent). Group 2 was identical to group 1, except that the comparison image depicted a man with low masculinity (100 percent). In group 3, participants saw two female faces: One of the faces was the unmodified original composite, and the comparison image had been modified to depict high masculinity (+100 percent). In group 4, the comparison image depicted a woman with low masculinity (100 percent). The survey website presented statements about men to groups 1 and 2, and statements about women to groups 3 and 4. Participants were asked to assign each statement to a leader with modified masculinity (high in groups 1 and 3, or low in groups 2 and 4), to a leader with unmodified masculinity or to neither leader. Because we were interested in attributions to the modified image featuring high or low masculinity, respectively, we coded such attributions for the purpose of analysis as 1 and other attributions (to the unmodified leader image or neither leader) as 0. Each participant was assigned 80 randomly drawn statements from the overall statement pool to reduce cognitive load. Participant gender was assessed as a control variable.

## Data Quality Check

Prior to the analyses, we checked the quality of the data. As in Study 1, we established that participants' responses would be excluded if (1) participants' response times were longer than 2 SD above the mean; (2)

participants had the same IP address; or (3) a participant's response pattern was indicative of a lack of attention to the questionnaire items (e.g., all statements were attributed to the same image). To keep the survey short, a separate question to test attention was not included. None of the participants met any of the criteria for exclusion.

## Analysis

The research design featured repeated answers from participants who each assessed multiple statements, so we used generalized estimating equations (GEE) to calculate regression parameters (Liang & Zeger, 1986). We were interested in whether participants attributed a statement to a particular image (1) or not (0), thus we specified a binary logistic model. GEE is an extension to generalized linear models that can process binary variables with repeated measurements (Hu, Goldberg, Hedeker, Flay, & Pentz, 1998). The estimated regression coefficients from the GEE method are similar to those from the standard logistic models, but in contrast to standard logistic models, GEE does not produce biased standard errors for data featuring repeated answers (Hu et al., 1998). An important consideration in GEE is the ratio of Level 1 data (in our case, the number of statements that the participants rated) to Level 2 data (in our case, the number of participants). If this ratio is too low,

Author's voice:  
Discuss Data Observations



**TABLE 5**  
**Study 2: Percentage of Respondents Assigning Competitive Statements to Women Leaders with Low and High Facial Masculinity**

Competitive Statements Concerning Women Leaders	Leader Masculinity	
	Low (%)	High (%)
1. If you have any weakness, she will take advantage of it	90	37
2. She is ruthless in business	87	46
3. When it comes to her competitors, she is even more cutthroat and nasty	86	51
4. She did whatever it took	83	44
5. Workers on the whole remarked how callous and cold she can be	78	50
6. Top managers who defy her openly tend to have short careers with Acme	77	47
7. Went behind people's backs and signed another deal	77	43
8. She was feared by those around her	76	67
9. Once a decision is made, it is final	75	40
10. She was less than kind to a few people to get where she is at	75	35
11. She wants it her way or you are out	75	56
12. There is only one boss, and that is her	74	60
13. She had revolutionary ideas and also stole a few ideas from competition	73	54
14. She will do whatever she can to take you over	73	49
15. She keeps the company held to a higher standard than others	73	56
16. Some peers fear open communication and respect her out of fear	71	68
17. She had to be the leader and they had to follow or they were shown the door	67	44
18. Employees are still fearful of losing their jobs because of her leadership style	67	40
19. She is not afraid to be a hard ass and many employees have found this out	67	50
20. You are fired!' is the way she resolves conflicts	67	53
21. Obliterated all who stood in her way	66	34
22. "Compromise is not part of her vocabulary," spoke a high level executive	65	61
23. Can push a bit hard	63	88
24. Engaged in cooperations with bigger rivals and used or bought smaller ones	60	31
25. Strives for win-win situation, but will stab you in the back	60	37
26. If you are not on board, you are dead to her	59	53
27. She fired everyone and hired all new employees in which she hand selected	58	45
28. A few jealous ones would talk bad about her behind her back	56	10
29. She came on and immediately terminated the quarreling management team	55	52
30. She would consider their opinions but she would have the final say	54	39
31. Making tough choices	40	58

*Note.* Participants were asked to evaluate a random subset of 80 out of 167 statements and were divided into two groups. Between 24 and 66 participants evaluated each statement for the woman low in masculinity (−100 percent). Between 24 and 64 participants evaluated each statement for the woman high in masculinity (+100 percent). For reasons of comparability, results are presented as percentages. Percentages indicate how many people agreed that a competitive statement was characteristic of the individual shown in the target image rather than the reference image or neither of the two. This list shows competitive statements which received more than 50 percent agreement.

GEE models fail to converge (Bell, Ferron, & Kromrey, 2008). The ratio for our dataset across the different conditions was on average 0.57 (SD = 0.04), and all models converged without errors. For each GEE regression model, we report the odds ratios (OR), the lower and upper limit of the confidence intervals, and the *p* values as indicators of significance (Cumming, 2014). The default value of OR = 1 shows the absence of an effect in OR. Values lower than 1 indicate a negative effect, and values greater than 1 indicate a positive effect.

## Results

To what extent are competitive statements attributed to men and women with low or high facial masculinity? Table 4 shows the results of the regression models, controlling for participant gender. The four regression models predict perceptions of competitiveness for

men (Models 1 and 2) and women (Models 3 and 4), with low facial masculinity (Models 1 and 3) and high facial masculinity (Models 2 and 4). First, we examined the results for men. Model 1 indicates that the chances for the man with low masculinity to be ascribed competitive attributes were low, with results tending toward significance (OR = 0.82 [0.66–1.02], *p* = .07), whereas Model 2 shows that the chances for the man with high masculinity to be attributed competitiveness were high (OR = 1.34 [1.04–1.73], *p* = .03).

Next, we examined the results for women. Women with both high and low masculinity were ascribed competitiveness. Model 3 shows that participants were likely to attribute competitive statements to the woman with low facial masculinity (OR = 3.39 [2.89–3.97], *p* < .01), and Model 4 indicates that the woman with high facial masculinity

TABLE 6

**Study 2: Percentage of Respondents Assigning Competitive Statements to Men Leaders with Low and High Facial Masculinity**

Competitive Statements Concerning Men leaders	Leader Masculinity	
	Low (%)	High (%)
1. Coworkers consider him very driven	33	64
2. Does not tolerate people trying to act like they are smarter or wiser than he is	27	63
3. Some employees think he is too harsh	31	63
4. He would handle conflicts with management with a strong will, unwilling to give in	33	63
5. He treats others with respect to a degree, but mostly believes he is right	31	61
6. Analyzing his competitors	32	61
7. Has high expectations	35	59
8. They would say he is always on top of them	33	59
9. It is only on rare occasions that he will listen to an alternate viewpoint	29	58
10. Known to walk out of meetings, red-faced, until he settles down	29	57
11. He was different than the rest. He wanted perfection	50	30
12. Academic excelling student	61	24
13. In college, he began to stand out because of his academic achievements	50	21
14. Was an excellent student through his high school and college years	57	21
15. He graduated with a 3.8 GPA	54	20
16. Graduated with honors	54	17
17. Two majors in college	61	13
18. Received high marks in school as well as excelling in his artistic pursuits	61	12

*Note.* Participants were asked to evaluate a random subset of 80 out of 126 statements and were divided into two groups. Between 40 and 56 participants evaluated each statement for the man low in masculinity (−100 percent). Between 40 and 56 participants evaluated each statement for the man high in masculinity (+100 percent). For reasons of comparability, results are presented as percentages. Percentages indicate how many people agreed that a competitive statement was characteristic of the individual shown in the target image rather than the reference image or neither of the two. This list shows competitive statements which received more than 50 percent agreement.

also had a high chance to be seen as competitive ( $OR = 3.41 [2.44-4.76], p < .01$ ).

Tables 5 and 6 list the competitive statements that more than 50 percent of participants assigned to the women and men leaders, respectively.

## Discussion

This study helps to illuminate people's preferences for particular leaders. Results indicate that those leaders who, in Study 1, were chosen more often to lead in competitive situations were considered to be more competitive in Study 2. Competitiveness was attributed to men with high facial masculinity, and to women at both extremes of masculinity (high *and* low). Men with low levels of masculinity, who had not been preferred in Study 1, were also not considered competitive in Study 2. These findings help explain people's leadership choices.

For the woman with high facial masculinity, the results corroborate prior research that found attributions of leadership (e.g., Hemmer & Kleiber, 1981). But for the woman with low masculinity, our findings are a surprising discovery. The woman with low facial masculinity is perceived to be competitive, and not cooperative as prior research suggests (Eagly & Karau, 2002).

Studies 1 and 2 examined people's leadership choices using hypothetical scenarios. In Study 3, we

asked whether our results concerning leader facial masculinity manifest in corporate field settings. Thus, in Study 3, we examined the facial features of men and women CEOs of large companies.

## STUDY 3

In this study, we showed images of men's or women's faces to research participants and asked them to rate those images for masculinity in comparison to a standard face of a man or a woman, respectively. Participants were not informed that the men's or women's faces they were rating were, in fact, the faces of CEOs of S&P 500 companies. We wondered: Would participants perceive differences in the facial masculinity of men CEOs and women CEOs?

## Participants and Procedure

We recruited 98 American adults via the online panel Microworkers to participate in a web-based research study. The sample was balanced in terms of gender (52 percent women) and quite diverse in terms of ethnicity (62 percent White, 16 percent Black, 8 percent Asian, 7 percent Latino, and 7 percent other). Participants averaged 35 years of age ( $SD = 12.03$ ), and 14 years ( $SD = 11.52$ ) of work



**TABLE 7**  
**Study 3: GEE Regression Results for Perceived Masculinity of Men and Women CEOs of S&P 500 Companies**

Perceived Masculinity of CEO <sup>a</sup>	Model 1		Model 2	
	Lower <sup>b</sup>		Higher <sup>b</sup>	
	Odds ratio	95% CI	Odds ratio	95% CI
Participant gender <sup>a</sup>	1.75**	1.43–2.14	1.79**	1.29–2.45
CEO gender <sup>a</sup>	1.65**	1.34–2.02	2.30**	1.64–3.23
Goodness of fit (QICC)	2,414.88		1,156.47	
Correlated data				
Number of participants	98			
Number of CEO images	50			
Number of observations	2,448			

*Note.* OR higher than one indicate a positive effect lower than 1 indicate a negative effect. The corrected quasi likelihood under Independence Model Criterion (QICC) for model fit is in the “small is better format.”

<sup>a</sup>Woman = 1, man = 0.

<sup>b</sup>Compared to the reference image.

\*\* $p < .01$

experience. After agreeing to take part in our research, participants accessed the survey website and were randomly assigned to two groups. One group ( $n = 50$ ; 48 percent women) was asked to assess the masculinity of men, and the other group ( $n = 48$ ; 56 percent women) was asked to assess the masculinity of women. We did not mix the two groups because we wanted to ensure that ratings of masculinity of men were independent of, and unaffected by, ratings of masculinity for women, and *vice versa*. Each group evaluated 25 images in total. The 25 images of men and the 25 pictures of women all featured CEOs of large companies<sup>6</sup>, but importantly, the study participants were not informed that they were rating CEOs.

We asked participants: “Please indicate how masculine each of the following men (women) is compared to the reference image.” Participants indicated their perception of each person’s masculinity compared to a reference image (1 = much less masculine, 2 = less masculine, 3 = just as masculine, 4 = more masculine, 5 = much more masculine). The reference image featured the man’s (woman’s) face unaltered in

masculinity that we had used in Studies 1 and 2 (Re et al., 2013). We used the reference image to calibrate participants’ perceptions of masculinity. The reference image was shown repeatedly throughout the questionnaire; thus, as participants scrolled down the page, they were still able to compare images.

### Data Quality Check

Prior to the analyses, we applied the same data quality checks as in Study 1. None of the participants were excluded.

### Results

We ran GEE regression models to assess whether differences between perceptions of masculinity among women and men CEOs were statistically significant. The models predicted ratings of “very low masculinity” and “very high masculinity” separately at first, and as a robustness check, we specified a multinomial model that included both categories simultaneously. For the purpose of these analyses, we coded perceptions of masculinity into two binary dependent variables: very low and very high masculinity. Very low masculinity was indicated by a score of 1 on the perception of masculinity scale, but not by scores from 2 to 5 on that scale. Thus we assigned a value of 1 for a score of 1, and a value of 0 for all other values (2–5). Likewise, very high masculinity was indicated by a score of 5 on the perception of masculinity scale, but not by scores from 1 to 4 on that scale. Thus we assigned a value of 1 for a score of 5, and a value of 0 for all other values (1–4). The use of binary dependent variables requires estimation by way of generalized linear models rather than by ordinary linear regression. As multiple raters evaluated the same pictures, the data are non-independent. Thus, as in Study 2, we ran regressions

<sup>6</sup> At the time this study was conducted, there were 25 women listed as CEOs of S&P 500 companies (Catalyst, 2014). All these women were included in our study. From the remaining 475 men listed as CEOs of S&P 500 companies, we randomly selected 25 to be included in our study. Specifically, we generated random numbers from 1 to 500 and then selected the CEOs of those companies whose rank in the S&P 500 list was identical to the random number. If, by coincidence, the CEO was a woman (one event) or if we were unable to obtain a verified image of the company CEO (two events), then we selected the CEO of the company listed subsequently on the ranking of the S&P 500 companies. All pictures were obtained through the Internet and formatted such that CEO’s faces were shown in roughly similar size.

using GEE, an extension of generalized linear models for repeated measurements.

Table 7 shows that more women CEOs than men CEOs were perceived to be much lower in masculinity than the reference picture (Model 1, OR = 1.65 [1.34–2.02],  $p < 0.01$ ) and, also, that more women CEOs than men CEOs were perceived to be much higher in masculinity than the reference picture (Model 2, OR = 2.30 [1.64–3.23],  $p < 0.01$ ). These findings suggest that among the 50 CEOs sampled from S&P 500 companies, there is greater range of masculinity among women CEOs than among men CEOs. More women CEOs than men CEOs were perceived to be at either end of the masculinity scale, thus having either high or low masculinity.

As a robustness check, we specified a multinomial GEE regression model that included both categories (i.e., “very low masculinity” and “very high masculinity”) together with the reference category. Thus we specified the dependent variable to have three possible values, corresponding to CEO masculinity being perceived as, first, low (1); second, medium (24); or, third, high (5). Results corroborated the findings that women CEOs more often than men CEOs were perceived to have very low and very high rather than medium masculinity (OR = 2.37 [1.91–2.94],  $p < .01$ ). Results remained statistically significant when participant gender was included as a control variable.

An inspection of the descriptive statistics corroborated the findings. Mean results indicated that on average, perceived masculinity of men and women CEOs was close to the middle of the scale (3), and that the mean ratings of masculinity in the group that compared men CEOs to the reference image ( $M = 2.92$ ,  $SD = 1.14$ ) were similar to the mean ratings of masculinity in the group that compared women CEOs to the reference image ( $M = 2.88$ ,  $SD = 0.97$ ). The distributions of facial masculinity ratings, however, differed in men and women CEOs, as indicated by a Kolmogorov–Smirnov test ( $p < .01$ ). Table 8 shows that among women CEOs, participants perceived there to be more extremes—in 21 percent of all cases, women

Author’s voice:

What was the most difficult or challenging aspect of this research project and paper?



CEOs were seen as being at the extreme ends of the scale (i.e., high [9 percent] or low [12 percent] in masculinity). Only in 11 percent of all cases were men CEOs at the extreme ends of the scale (i.e., high [4 percent] or low [7 percent] in masculinity).

## Discussion

This study investigated whether in a field context, distributions of facial masculinity differ between men and women leaders. Among women CEOs, we find a comparatively greater range of facial masculinity than among men CEOs of S&P 500 companies. Such companies face considerable competition (Friedman, 2005; Gupta, Govindarajan, & Wang, 2008) and Studies 1 and 2 suggest that in such contexts, women with either low or high levels of masculinity would be preferred. In line with this conclusion, the results of this study suggest that not only women with high masculinity but also women with low masculinity become leaders in the corporate world. This finding challenges existing theory and research, which posits a universal advantage of facial masculinity in competitive settings (Re et al., 2013; Spisak, Dekker, et al., 2012).

Somewhat surprisingly, in this study, men CEOs were not perceived to be particularly high in facial masculinity. Based on existing research and Studies 1 and 2, we were expecting people to perceive male CEOs to feature high rather than medium masculinity. But ratings suggested that many CEOs were perceived relatively similarly to the reference image depicting the unmodified original composite man. One possible explanation is that the man shown in the reference image, in the absence of other comparisons, was perceived as quite masculine.

**TABLE 8**  
**Study 3: Ratings of Masculinity for Men and Women CEOs of S&P 500 Companies**

Rating <sup>a</sup>	Frequencies		Percentages	
	Woman CEOs	Man CEOs	Woman CEOs	Man CEOs
1 (Much less masculine)	139	92	12	7
2 (Less masculine)	308	340	26	27
3 (Similarly masculine)	370	498	31	40
4 (More masculine)	276	266	23	21
5 (Much more masculine)	107	54	9	4
Number of CEO images	25	25		
Number of participants	48	50		

<sup>a</sup>Compared to the reference image.

Author's voice:

Was there anything that surprised you about the findings?



This points to the limitation that the images we used across studies allow inferences concerning *relative* deviations in masculinity from the baseline original composite, but the *absolute* masculinity of the baseline original composite is difficult to discern. Therefore, findings from Study 3 cannot be interpreted as evidence that men CEOs are of average masculinity; rather, that men CEOs (in contrast to women CEOs) were perceived more similarly in terms of their facial masculinity.

## GENERAL DISCUSSION

In a series of three studies, we found that leadership choices based on facial masculinity differ for women and men leaders. Findings from Study 1 indicate that masculine-looking men and women are more preferred in competitive than in collaborative contexts. However, we also discovered that women with low facial masculinity are chosen more often as leaders in competitive than in collaborative scenarios. This is remarkable because previous research has linked low facial masculinity to perceptions of warmth and cooperativeness (Perrett et al., 1998). Study 2 corroborated that, indeed, women with low facial masculinity are attributed competitiveness, whereas men with low facial masculinity are not seen as competitive. Given these findings we expected that in general business settings, we would find men leaders who look rather masculine and women leaders who look either rather masculine or not masculine at all. This notion was supported by Study 3, in which we found a greater range of facial masculinity among women CEOs than among men CEOs of S&P 500 companies.

This research confirms the basic assumption of biosocial models of leadership: People's perceptions of facial masculinity systematically affect their choices of leaders and the attributes that they assign to leaders. But our findings indicate that leadership choices for men and women differ more than is suggested by prior studies (Re et al., 2013; Spisak, Dekker, et al., 2012). Extant theory and research converge on the notion that

both women and men benefit in leadership selection processes from facial masculinity, particularly in competitive settings, but our findings suggest that women can also benefit from low facial masculinity.

Perhaps the most striking discovery of our research was that women with low facial masculinity were ascribed competitiveness. These attributions go beyond the typical gender-biased leadership prototype (Hogue & Lord, 2007). Gender stereotype research suggests that low facial masculinity should facilitate schemas concerning stereotypically female qualities, such as communality, warmth, and solidarity (Zebrowitz & Montepare, 2008). Indeed, it has been argued that followers expect feminine-looking leaders to be cooperative (Spisak, Homan, et al., 2012). But competitiveness among women has long been ignored or overlooked (Clutton-Brock & Huchard, 2013). Our research shows that low facial masculinity in women is linked to competitiveness and not only to cooperativeness as prior research suggests.

Biosocial models of leadership posit that the linkage between facial masculinity and leadership choice is rooted in hormonal processes that shape both facial appearance and behavior (Spisak, Dekker, et al., 2012). While for men the link between testosterone, facial masculinity, and status-striving behavior seems clear (Stanton & Schultheiss, 2009), for women, there is some evidence that dominant behavior is not only due to high levels of testosterone, but is also associated with a potent form of estrogen (Stanton & Schultheiss, 2007, 2009). Facial masculinity results from the interplay of testosterone and estrogen. Relative changes in visual cues indicative of testosterone and estrogen levels are reflected in the facial composites used in this study (Perrett et al., 1998; Re et al., 2013). To perceivers, the image that we used of the woman's face with low facial masculinity likely signaled relatively higher levels of estrogen (Perrett et al., 1998; Smith et al., 2006). If estrogen makes faces less masculine, but is also linked to dominant behavior (Stanton & Schultheiss, 2007, 2009), then our discoveries that women with low facial masculinity are chosen as leaders and seen as competitive appear quite plausible and in line with recent biological research on hormones. Thus striving for dominance is associated with the hormone testosterone in men, but appears to be linked with the hormone estrogen in women (Schultheiss, 2013).

Our findings in turn suggest that evolutionary-based implicit leadership theories are more sophisticated than previously assumed. Apparently, people do not only take manifestations of testosterone into account when conferring leadership on others, but also manifestations of estrogen. Our research suggests that not only are women with low facial masculinity suitable leaders in communal, cooperative settings

Author's voice:

If you were able to do this study again, what if anything would you do differently?



involving a “niche for feminine leadership” (Re et al., 2013; Spisak, Dekker, et al., 2012: 6), but that they are also seen (Study 1) and employed (Study 3) as leaders in competitive settings. The leadership disadvantages arising from low facial masculinity appear to affect only one gender: men. In line with previous research that suggests that men’s status suffers to the extent that they do not fit with gendered expectations (Moss-Racusin et al., 2010), men with low facial masculinity were not seen as competitive leaders in our research.

This study is an example of how a more differentiated understanding of social and organizational issues—such as the question of who should be the leader—can be advanced through a more fine-grained understanding of the complex biological, and particularly hormonal, processes that underpin overt appearance and behavior (Becker & Menges, 2013). The effects revealed in this research suggest patterns that are much more nuanced than those derived from a binary viewpoint concerning biological sex. Thus, these effects challenge extant gender stereotypes and implicit leadership research that has, for the most part, neglected to take biologically rooted differences *within* sex categories into account. Future theory and research should move from the examination of differences between men and women to the consideration of differences among men and differences among women.

Our research is limited in its focus on one cultural setting and the use of a relatively restricted set of images. Prior research shows that facial masculinity effects manifest across cultures (Re et al., 2013; Spisak, Dekker, et al., 2012), so we assume the effects would hold outside North America. Nonetheless, we encourage crosscultural research to replicate and extend on our findings. The images used in this research were validated in previous research (Re et al., 2013) and were suited to our research context. Future research should aim to create and utilize images that allow for inferences concerning absolute masculinity. Such images would make it possible to resolve the question of why a large proportion of men CEOs’ in Study 3 were perceived as similar in masculinity to the reference image, even though most research points out that men leaders benefit from high facial masculinity in competitive settings.

Finally, future research should investigate how low facial masculinity among men and women leaders relates to perceived femininity. There may be different types of perceived femininity, signaling either competitiveness (Stanton & Schultheiss, 2007) or cooperativeness (Perrett et al., 1998).

In conclusion, this research suggests that people seem to implicitly understand facial markers of the hormones testosterone and estrogen as signs for

competitiveness and leadership suitability in men and women with high facial masculinity, as well as in women with low facial masculinity. What people read in the faces of Yahoo’s Marissa Mayer, Hewlett Packard’s Meg Whitman, and Facebook’s Sheryl Sandberg is that these nonmasculine-looking women are perceived as competitive, and therefore well suited to lead their businesses toward the achievement of organizational goals and future success.

## REFERENCES

- Baldwin, M. W. 1992. Relational schemas and the processing of social information. *Psychological Bulletin*, 112: 461–484.
- Becker, W. J., & Menges, J. I. 2013. Biological implicit measures in HRM and OB: A question of how not if. *Human Resource Management Review*, 23: 219–228.
- Bell, B.A., Ferron, J.M., & Kromrey, J.D.. 2008. Cluster size in multilevel models: The impact of sparse data structures on point and interval estimates in two-level models. *American Statistical Association, Joint Statistical Meetings, Proceedings of the Section on Survey Research Methods*. Alexandria, VA: American Statistical Association.
- Boothroyd, L. G., & Perrett, D. I. 2006. Facial and bodily correlates of family background. *Proceedings, Biological Sciences*, 273: 2375–2380.
- Bohannon, J. 2011. Social Science for Pennies. *Science*, 334: 307.
- Brands, R. A., Menges, J., & Kilduff, M. 2015. The leader in social network schema: Perceptions of network structure affect gendered attributions of charisma. *Organization Science*, 4: 1210–1225.
- Buhrmester, M., Kwang, T., & Gosling, S. D. 2011. Amazon’s mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6(1): 3–5.
- Catalyst. Women CEOs of the S&P 500. <http://www.catalyst.org/knowledge/women-ceos-sp-500>. Accessed September 15, 2014.
- Chafe, W. L. 1980. *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*. Norwood, MA: Ablex Pub. Corp.
- Clutton-Brock, T. H., & Huchard, E. 2013. Social competition and selection in males and females. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 368: 1–15.
- Cumming, G. 2014. The new statistics: why and how. *Psychological Science*, 25(1): 7–29.



- De Waal, F. 1982. *Chimpanzee politics: Sex and power among apes*. London, UK: Jonathan Cape. Baltimore, MD: Johns Hopkins University Press.
- Eagly, A. H., & Karau, S. J. 2002. Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109: 573–598.
- Feinberg, D. R., & Jones, B. C. 2005. The voice and face of woman: One ornament that signals quality? *Evolution and Human Behavior*, 26: 398–408.
- Friedman, T. 2005. *The world is flat: A brief history of the globalized world in the 21st century*. Farrar:488. New York: Farrar, Straus & Giroux.
- Gupta, A. K., Govindarajan, V., & Wang, H. 2008. *The quest for global dominance: Transforming global presence into global competitive advantage*. San Francisco, CA: John Wiley & Sons.
- Hehman, E., Carpinella, C. M., Johnson, K. L., Leitner, J. B., & Freeman, J. B. 2014. Early processing of gendered facial cues predicts the electoral success of female politicians. *Social Psychological & Personality Science*, 5: 815–824.
- Hemmer, J. D., & Kleiber, D. A. 1981. Tomboys and sissies: Androgynous children? *Sex Roles*, 7: 1205–1212.
- Hinkin, T. R. 1995. A review of scale development practices in the study of organizations. *Journal of Management*, 21: 967–988.
- Hogue, M., & Lord, R. G. 2007. A multilevel, complexity theory approach to understanding gender bias in leadership. *The Leadership Quarterly*, 18: 370–390.
- Hu, F. B., Goldberg, J., Hedeker, D., Flay, B. R., & Pentz, M. A. 1998. Comparison of population-averaged and subject-specific approaches for analyzing repeated binary outcomes. *American Journal of Epidemiology*, 147: 694–703.
- Jiménez, M., Aguilar, R., & Alvero-Cruz, J. R. 2012. Effects of victory and defeat on testosterone and cortisol response to competition: Evidence for same response patterns in men and women. *Psychoneuroendocrinology*, 37: 1577–1581.
- Jæger, M. M. 2011. “A Thing of Beauty is a Joy Forever”? Returns to physical attractiveness over the life course. *Social Forces; a Scientific Medium of Social Study and Interpretation*, 89(3): 983–1003.
- Lefevre, C. E., Lewis, G. J., Perrett, D. I., & Penke, L. 2013. Telling facial metrics: Facial width is associated with testosterone levels in men. *Evolution and Human Behavior*, 34: 273–279.
- Liang, K., & Zeger, S. 1986. Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1): 13–22.
- Little, A. C. 2014. Facial appearance and leader choice in different contexts: Evidence for task contingent selection based on implicit and learned face-behaviour/face-ability associations. *The Leadership Quarterly*, 25: 865–874.
- Little, A. C., Burriss, R. P., Jones, B. C., & Roberts, S. C. 2007. Facial appearance affects voting decisions. *Evolution and Human Behavior*, 28: 18–27.
- Lord, R. G., & Dinh, J. E. 2014. What have we learned that is critical in understanding leadership perceptions and leader-performance relations? *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 7: 158–177.
- Lord, R. G., Foti, R. J., & De Vader, C. L. 1984. A test of leadership categorization theory: Internal structure, information processing, and leadership perceptions. *Organizational Behavior and Human Performance*, 34: 343–378.
- Lutchmaya, S., BaronCohen, S., Raggatt, P., Knickmeyer, R., & Manning, J. T. 2004. 2nd to 4th digit ratios, fetal testosterone and estradiol. *Early Human Development*, 77: 23–28.
- Moss-Racusin, C. A., Phelan, J. E., & Rudman, L. A. 2010. When men break the gender rules: Status incongruity and backlash against modest men. *Psychology of Men & Masculinity*, 11(2): 140–151.
- Murray, G. R., & Schmitz, J. D. 2011. Caveman politics: Evolutionary leadership preferences and physical stature. *Social Science Quarterly*, 92: 1215–1235.
- Neisser, U. 1976. *Cognition and reality: Principles and implications of cognitive psychology*. New York, NY: W H Freeman/Times Books/Henry Holt & Co.
- Newport, F., & Wilke, J. 2013. Americans still prefer a male boss. *Gallup*. <http://www.gallup.com/poll/165791/americansprefermaleboss.aspx>. Accessed September 10, 2014.
- Penton-Voak, I. S., Pound, N., Little, A. C., & Perrett, D. I. 2006. Personality judgments from natural and composite facial images: More evidence for a “kernel of truth” in social perception. *Social Cognition*, 24(5): 607–640.
- Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., Henzi, S. P., Castles, D. L., & Akamatsu, S. 1998. Effects of sexual dimorphism on facial attractiveness. *Nature*, 394: 884–887.
- Pound, N., Penton-Voak, I. S., & Surridge, A. K. 2009. Testosterone responses to competition in men are related to facial masculinity. *Proceedings, Biological Sciences*, 276: 153–159.
- Re, D. E., DeBruine, L. M., Jones, B. C., & Perrett, D. I. 2013. Facial cues to perceived height influence leadership choices in simulated war and peace contexts. *Evolutionary Psychology*, 11: 89–103.



- Rhodes, G. 2006. The evolutionary psychology of facial beauty. *Annual Review of Psychology*, 57: 199–226.
- Schein, V. E. 1973. The relationship between sex role stereotypes and requisite management characteristics. *The Journal of Applied Psychology*, 57: 95–100.
- Schein, V. E. 2001. A global look at psychological barriers to women's progress in management. *The Journal of Social Issues*, 57: 675–688.
- Schultheiss, O. C. 2013. The hormonal correlates of implicit motives. *Social and Personality Psychology Compass*, 7: 52–65.
- Smith, M. J. L., Perrett, D. I., Jones, B. C., Cornwell, R. E., Moore, F. R., Feinberg, D. R., Boothroyd, L. G., Durrani, S. J., Stirrat, M. R., Whiten, S., Pitman, R. M., & Hillier, S. G. 2006. Facial appearance is a cue to oestrogen levels in women. *Proceedings, Biological Sciences*, 273: 135–140.
- Spisak, B. R., Dekker, P. H., Krüger, M., & van Vugt, M. 2012. Warriors and peacekeepers: Testing a biosocial implicit leadership hypothesis of intergroup relations using masculine and feminine faces. *PLoS One*, 7: e30399.
- Spisak, B. R., Homan, A. C., Grabo, A., & Van Vugt, M. 2012. Facing the situation: Testing a biosocial contingency model of leadership in intergroup relations using masculine and feminine faces. *The Leadership Quarterly*, 23: 273–280.
- Stanton, S. J., & Schultheiss, O. C. 2007. Basal and dynamic relationships between implicit power motivation and estradiol in women. *Hormones and Behavior*, 52: 571–580.
- Stanton, S. J., & Schultheiss, O. C. 2009. The hormonal correlates of implicit power motivation. *Journal of Research in Personality*, 43: 1–15.
- Thornhill, R., & Gangestad, S. W. 1999. Facial attractiveness. *Trends in Cognitive Sciences*, 3: 452–460.
- Willis, J., & Todorov, A. 2006. First impressions: Making up your mind after a 100ms exposure to a face. *Psychological Science*, 17: 592–598.
- Zebrowitz, L. A., & Montepare, J. M. 2008. Social psychological face perception: Why appearance matters. *Social and Personality Psychology Compass*, 2: 1497.



**Raphael Silberzahn** is a research associate at the University of Cambridge - Judge Business School, from where he also obtained his PhD.

**Jochen Menges** is a University Professor of Leadership and Human Resource Management at the WHU - Otto Beisheim School of Management in Düsseldorf, Germany. He obtained his PhD from the University of St. Gallen in Switzerland.

